A "Virtual Unix" for Tops-20

Available from:

Jay Lepreau
Computer Science Dept.
Univ. of Utah
Salt Lake City, UT
84112

Lepreau@UTAH-20
801-581-8332
-581-8224

decvax!{harpo,randvax}!UTAH-cs!Lepreau

WHAT'S INVOLVED IN THIS PROJECT?

① C COMPILER

- PORTABLE C COMPILER (PCC)
- AT LEAST 1/2 THE WORK
- BOOTSTRAPPING

② PROVIDING THE V7 UNIX ENVIRONMENT

FIRST C LIBRARY ROUTINES + SOME SYSTEM CALLS AS LINKABLE OBJECT MODULES

THEN ⇓

PSEUDO-KERNEL PROVIDING ALMOST ALL UNIX SECTION 2 SYSTEM CALLS (+ SOME SEC. 3)

(3 - PORTING USER PROGRAMS)

TRIVIAL - MAYBE

~

WHY DO THIS ANYWAY?

- SYSTEMS LANG. ON 20
- COMMON SYSTEMS. LANG.
  BCPL – C
- EXTENDED. ADDRESSING SUPPORT ~1 MB
- PORT UNIX SOFTWARE

(THIS GREW ON US)

WHY NOT JUST UNIX?

- BOTH TOPS-20 & UNIX ARE "NICE"

TOPS-20:

- COMND JSYS
- GENERAL PAGE-MAPPING
- FRONT-END/BACKEND
- FILE VERSION #'S.

UNIX:

- MOTHERHOOD - WHAT WE KNOW & LOVE
- SHELL, CSH . REDIRECTION
  PIPES
  PROGRAM.
- UTILS IN C! MAKE, GREP, YACC

GOALS

- SAME FILE FMT FOR TOPS-20 + "VUNIX"
- ACCESS TO BOTH ENVIRONMENTS FROM C
- SAME LOAD MODULES BOTH W/ + W/O THE SHELL (WITH SOME SHELL FEATURES ANYWAY)
- KEEP LOAD MODULES SMALL
- "KERNEL" CHANGES INVISIBLE TO USER PGMS.
- <u>NO</u> HACKING OF TOPS-20 (MACRO - YECH!)

BOOTSTRAPPING HOSTS

1. 11/34 V6
   - pcc + yacc V7 split I/D

2. 11/45 V6

3. 11/45 V7

C COMPILER ISSUES

- PDP-10 QUICKIE:
  - WORD ADDRESSABLE (PCC PROB)
  - 36 BIT WORD, ANY SIZE BYTE
  - STD. BYTE SIZE == 7

    36 % 7 != Ø
- C "SPEC" PROBLEMS
  - SIZEOF IN BYTES

- PCC DOESN'T LIKE IT
- WORST: LIBRARIES ASSUME CHARS GET EVERYTHING

⇒ 9 BIT BYTES INTERN

MAPPED TO 7 EXTERN

SAM LEFFLER'S WORK AT CASE
- WORD ADDR.
- DETAILED DOC

FIRST TEST PROGRAM:

```
int a;
foo()
{
a = 0;
}
```

```
int a;
foo()
{
 a = 0;
}
```

```
c9 line 4: compiler error: expression causes compiler loop: try simplifying
```

COMPILER ERROR: EXPR. CAUSES
COMPILER LOOP: TRY SIMPLIFYING. !

THEN I HACKED UNIX FOR 8 MOS.

LIBRARY

- USUALLY JUST DO STDIO TO SUPPORT C COMPILER
  LOTS OF PGMS DON'T USE IT –

- SO JUST: IMPLEMENT:
  OPEN, CLOSE, READ, WRITE
  MALLOC (SBRK), FREE, ISATTY
  DOPRNT

  AND GET IT ALL!

- WHY WRITE IT WHEN YOU CAN "BORROW" IT?

NEXT - "VIRTUAL" UNIX

- RTS IN LOCORE, INCLUDES CMD LINE PARSING

- USER PGM CHECKS FLAG IN LOCORE, MAYBE LOADS RTS.

- NEW PROCESS:
  - RTS IN SHARED MEMORY
  - PROC & USER STRUCTS, ENV.
  -
  - PIPES.

- BORROW KERNEL CODE TOO!
  - UFALLOC, FALLOC TO ALLOC. FD'S & FILES
  - PIPE.C - EMULATE WAKEUP & SLEEP <u>EXCEPT</u> UNIX USES INTERRUPT MASKING

UNIX + TOPS-20

- SIMILARITIES
  - FILES ARE BYTE-STREAMS
  - FILESYS IS A TREE
  - NOTION OF "CURRENT DIRECTORY" (DIFFERENT SEMANTICS)

- DIFFERENCES
  - LINE TERMINATOR CR/LF - LF
    MUNDANE BUT A BITCH! FSEEK
  - STRICT HIERARCHY OF PROCESSES
    PID'S ARE RELATIVE
  - LONG FILENAMES
  - NO LINKS
  - "CHDIR" SEMANTICS
  - DIR QUOTAS

⇒ EASIER TO PORT FILESYS + PROC STRUCTURE

*** SECTION 3 I/O ***

fopen, freopen, fdopen
fclose, fflush
fread, fwrite
printf, fprintf, sprintf
scanf, fscanf, sscanf
gets, fgets
puts, fputs
setbuf
isatty
ungetc
getc, fgetc, getchar, getw
putc, fputc, putchar, putw
clearerr, feof, ferror, fileno

CURRENT ROUTINES 1/22/81

ALL I/O

*** SECTION 3 STRING ROUTINES ***

index, rindex, strcat, strcmp, strcpy, strlen, strncat, strncmp, strncpy

*** SECTION 3 CTYPE MACROS ***

isalpha, isupper, islower, isdigit, isxdigit, isspace, ispunct
isalnum, isprint, iscntrl, isascii, toupper, tolower, toascii

*** SECTION 3 MISC ***

malloc, calloc, realloc, valloc, free
atoi, atol, atof
setjmp, longjmp
ctime, timezone
perror
qsort
mktemp
getenv
re_comp, re_exec
crypt, encrypt, setkey
nargs

*** SECTION 3 MATH ***

sin, cos
modf, frexp
ecvt, fcvt

*** SECTION 2 SYSTEM CALLS ***

open, close
read, write
ioctl
stat, fstat
execl, execv
fork, wait
brk, sbrk
exit
getuid
access
chdir
unlink
abort

SYSTEM CALLS

*** USE WITH CAUTION DUE TO CR/LF PROBLEM ***

lseek, tell
fseek, ftell, rewind

*** TOPS-20 ROUTINES ***

fnstd, fnunix
jsys ~120 LINES OF ASM ONLY

UNIX SOFTWARE RUNNING NOW

EGREP

GREP

LEX

YACC

SED

COMM

CB

DC

OD

UNIQ

WC

(CPP, PCC)

RAIN

WORMS

TERMCAP

WHERE'S ROGUE?

PORTABILITY PROBLEMS IN RETROSPECT

LESSONS FOR PGMERS!

- "EXTERN" DECLARATION
- LONG EXTERNAL NAMES
- CASE-SENSITIVE " "
- CHAR C; C = GETCHAR();
  - INEFFICIENT
  - WRONG
- <VARARGS.H>
- SEEK

- STRONGLY-TYPED LANGUAGE ESSENTIAL!
- CPP <u>IMPORTANT</u> FOR X-COMPILING.

LESSONS LEARNED - CLICHÉS

- DON'T WASTE TIME IN INADEQAUE ENVIRONMENT — WAIT OR QUIT

- GET LUCKY (LEFFLER)

- BOOTSTRAPPING:

  TIME TO COMPLETE $\approx e^{\text{FTP-TIME}}$

- DON'T WORRY ABOUT DETAILS OF C SEMANTICS AT BEGINNING (36 >> 32)

- **DO** SPEND TIME WORKING OUT THE CALLING CONVENTION
  - VARARGS
  - DEBUGGING INFO - EXTRA

- PCC ITSELF — WITH 16 REGS

  DON'T WORRY ABOUT CONSERVATION OF REGISTERS!